

October 28, 2024

**VIA ECF**

Hon. Ona T. Wang  
United States District Court  
Southern District of New York

Re: *Authors Guild et al. v. OpenAI, Inc., et al.*, Case Nos. 1:23-cv-08292-SHS & 1:23-cv-10211-SHS [Dkt. 235–Plaintiffs’ Letter Motion Seeking Additional Custodians]

Dear Magistrate Judge Wang:

Over the past several months, the Authors Guild plaintiffs have sought invasive and burdensome custodial discovery from OpenAI that is untethered to their copyright infringement claims and disproportionate to the needs of the case. Initially, plaintiffs demanded 30 custodians. After OpenAI agreed to several of those, plaintiffs moved the goalposts and demanded others. Then plaintiffs changed their minds, dropped several of *those* custodians, and demanded new ones. And so on. Plaintiffs’ demands and justifications for new custodians shift by the day. The only constant has been the demand for “more,” regardless of whether the custodians “would provide unique, relevant and noncumulative evidence.” *See Coventry Cap. US LLC v. EEA Life Settlements, Inc.*, No. 17-cv-07417, 2021 WL 961750, at \*2 (S.D.N.Y. Mar. 15, 2021).

The Court’s guidance at the September 12 hearing should have put an end to this: it ordered plaintiffs to review OpenAI’s productions *before* demanding more. 9/12 Hr’g. Tr. at 101:20-23. That approach remains sensible. OpenAI has already agreed to produce documents from 24 custodians, including its most senior executives, and is in the midst of rolling out substantial productions, including over 6,000 documents produced this month. But instead of following the Court’s guidance, plaintiffs have flouted it. They did not bother to review OpenAI’s productions before burdening the Court with their demand for “more.” In fact, plaintiffs’ motion renews their demand for the same 6 custodians at issue in their prior motion – and adds two new ones – relying exclusively on documents that OpenAI produced in May and June. In other words, no new information has come to light that justifies plaintiffs’ “renewed” motion; instead, it rehashes the same failed arguments as the last one. Plaintiffs have disregarded the Court’s guidance and resumed the drumbeat for duplicative, disproportionate discovery before reviewing what they have. These scorched earth tactics should not be condoned, and their motion should be denied.

**I. Plaintiffs have disregarded the Court’s guidance.**

OpenAI has produced documents from 24 custodians, including its CEO, President, former CTO, and key members, including senior leadership, of its technical, research, product and business teams. OpenAI’s productions include documents regarding the technical design, development, and operation of ChatGPT; the training of OpenAI’s models; compliance and safety issues; and

OpenAI's partnerships and agreements. By the end of this month, OpenAI will have produced and/or made available for inspection nearly 49,000 documents and 175 TB of data. Those productions are ongoing.

Notwithstanding these efforts, plaintiffs move to compel additional custodians, without regard for whether those individuals have unique, non-cumulative information. This is precisely what the Court counseled against at the September 12 hearing when it pressed plaintiffs to explain, "how is it that you know that these six custodians are not cumulative? [That] [t]he documents and information from these six additional proposed custodians would not be cumulative or duplicative of what's being produced?" 9/12 Hr'g. Tr. at 33:8-11. Ultimately, the Court directed plaintiffs to review "*the production that you're getting from the custodians already agreed to and then see if there is a need to further search those . . . additional custodians.*" *Id.* at 101:20-23 (discussing as to both *Authors Guild/Alter* and *New York Times* cases) (emphasis added). Plaintiffs have not done so, and their motion fails on this basis alone.

## II. Plaintiffs have not met their burden to justify adding new custodians.

Plaintiffs' motion also fails because it is their burden to "demonstrate that the additional requested custodians would provide *unique* relevant information not already obtained." *Fort Worth Emps. Ret. Fund v. J.P. Morgan Chase & Co.*, 297 F.R.D. 99, 107 (S.D.N.Y. 2013). They have not met it here.

**Chelsea Voss (1) & Jan Leike (2)**: Plaintiffs requested Ms. Voss and Mr. Leike as custodians on October 14. The parties conferred on October 21, and OpenAI informed plaintiffs that it would consider and respond to their request by October 28, the deadline set by the ESI order. *See* Dkt. 201 at 3. Rather than wait for OpenAI's response, plaintiffs filed their motion on October 23 – five days before OpenAI's response was due. In their haste to file, plaintiffs' motion makes a number of misstatements. For example, plaintiffs claim that "[o]nly Chelsea Voss has complete information" about certain issues, Dkt. 234 at 2, but then refer to documents showing other custodians have that same information.<sup>1</sup> In fact, all but one of the communications plaintiffs cite are with designated custodians, including Ms. Voss's manager and team members. The remaining document, Dkt. 193-5, shows that multiple designated custodians were aware of and tracked Voss's work on ebook data; it does not show that she has unique knowledge.

Plaintiffs demand Mr. Leike on the grounds that "OpenAI has not produced a single custodian in the [REDACTED] channel." Dkt. 234 at 2. This is not true. OpenAI's designated custodians already include [REDACTED]. Plaintiffs' other justification—that Mr. Leike "disagree[d] with OpenAI leadership about the company's core priorities"—is unavailing and irrelevant. *Id.* at 2. Plaintiffs' copyright claims do not warrant unlimited custodial discovery into any OpenAI employee who might hold such views.

**Ilya Sutskever (3)**: Plaintiffs' justifications for adding Mr. Sutskever are ever-shifting. *See* Dkt. 193 at 2. Plaintiffs' latest argument is that Mr. Sutskever talked to outside investors and "would have been central in decisions about the use of data to train models and the use of tools to restrict the output of models." Dkt. 234 at 2. But none of those topics is unique to Mr. Sutskever. OpenAI's 24 designated custodians include multiple individuals knowledgeable on those

<sup>1</sup> Dkt. 234, Exs. 2 ([REDACTED]), 4 ([REDACTED]), 11 (same), 20 (same), 24 ([REDACTED]).

subjects, including its CEO, President, former CTO, head of its safety team, as well as *nine* individuals who worked on training the relevant models. Plaintiffs' claim that Mr. Sutskever has "unique files" is likewise unsupported. *Id.* They fail to articulate the relevance of [REDACTED] (*id.*, Ex. 23); or explain why Mr. Sutskever's [REDACTED] (*id.*, Ex. 21) justifies a full search of his files. Plaintiffs also cite a 2-line exchange between Mr. Sutskever and another employee (*id.*, Ex. 17), but do not explain why it shows that he has unique information.

**Andrew Mayne (4):** Plaintiffs justify their demand for Mr. Mayne's files on the grounds that he "is OpenAI's only [REDACTED] [REDACTED]" Dkt. 234 at 2, but that was not his role. In any event, plaintiffs fail to explain how "[REDACTED]" are relevant. *Id.* Nor did Mr. Mayne, who was subpoenaed by plaintiffs, direct them to seek documents from OpenAI in response; instead, he agreed to search for many categories of documents and found none. *Id.* at 3, Ex. 26. Plaintiffs' disappointment does not justify a fishing expedition.

**Shantanu Jain (5):** Plaintiffs' sole justification for Mr. Jain is his alleged involvement in investigating [REDACTED] for model training. But OpenAI has designated nine custodians who worked on model training, including [REDACTED]. Dkt. 198 at 2-3.

**Qiming Yuan (6):** Plaintiffs' claims about Mr. Yuan's "unique" knowledge are unsupported by the documents, which confirm that at least two other custodians were aware of his work. Dkt. 234, Exs. 2, 5. And Mr. Yuan's supervisor, and his supervisor's supervisor, are also custodians.

**Jong Wook Kim (7):** Plaintiffs have abandoned their initial justification for Mr. Kim, *see* Dkt. 193 at 3, and concocted a new one—[REDACTED]. But the exhibits plaintiffs cite confirm [REDACTED]. *See* Dkt. 193-5 at 14; Dkt. 234, Ex. 18. And the Slack channel in Exhibit 7 [REDACTED]; it does not show Mr. Kim has unique information.

**Cullen O'Keefe (8):** Plaintiffs rely exclusively on the [REDACTED] cited in their original motion, Dkt. 193-5, but inexplicably dismiss as "beside the point" that Mr. O'Keefe sourced the information in the document from three custodians. Dkt. 234 at 3. And plaintiffs' preference to have a custodian on undefined "governance" issues, *id.*, is insufficient to justify a burdensome privilege review of an attorney's files. Their motion should be denied.

Sincerely,

KEKER, VAN NEST &  
PETERS LLP

LATHAM & WATKINS LLP

MORRISON FOERSTER

/s/ Michelle Ybarra  
Michelle Ybarra

/s/ Elena Nightingale Dawson  
Elena Nightingale Dawson

/s/ Carolyn Homer  
Carolyn Homer